Alzheimer's & Dementia 14 (2018) 1580-1588

Alzheimer's & Dementia

Featured Article

Elevated DNA methylation across a 48-kb region spanning the *HOXA* gene cluster is associated with Alzheimer's disease neuropathology

Rebecca G. Smith[a], Eilis Hannon[a], Philip L. De Jager[b,c,d], Lori Chibnik[b,c], Simon J. Lott[e], Daniel Condliffe[f], Adam R. Smith[a], Vahram Haroutunian[g,h,i], Claire Troakes[e], Safa Al-Sarraj[e], David A. Bennett[j], John Powell[e], Simon Lovestone[k], Leonard Schalkwyk[l], Jonathan Mill[a,*,1], Katie Lunnon[a,**,1]

[a]*Institute of Clinical and Biomedical Science, University of Exeter Medical School, RILD Building, Royal Devon & Exeter Hospital Campus, Exeter, Devon, UK*
[b]*Program in Translational NeuroPsychiatric Genomics, Departments of Neurology and Psychiatry, Institute for the Neurosciences, Brigham and Women's Hospital, Boston, MA, USA*
[c]*Harvard Medical School, Boston, MA, USA*
[d]*Department of Neurology, Columbia University College of Physicians and Surgeons, Columbia University Medical Center, New York, NY, USA*
[e]*Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK*
[f]*Queen Mary University of London, London, UK*
[g]*Department of Psychiatry, The Icahn School of Medicine at Mount Sinai, New York, NY, USA*
[h]*Department of Neuroscience, The Icahn School of Medicine at Mount Sinai, New York, NY, USA*
[i]*JJ Peters VA Medical Center, Bronx, NY, USA*
[j]*Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA*
[k]*Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK*
[l]*University of Essex, Colchester, UK*

Abstract

Introduction: Alzheimer's disease is a neurodegenerative disorder that is hypothesized to involve epigenetic dysregulation of gene expression in the brain.
Methods: We performed an epigenome-wide association study to identify differential DNA methylation associated with neuropathology in prefrontal cortex and superior temporal gyrus samples from 147 individuals, replicating our findings in two independent data sets (N = 117 and 740).
Results: We identify elevated DNA methylation associated with neuropathology across a 48-kb region spanning 208 CpG sites within the *HOXA* gene cluster. A meta-analysis of the top-ranked probe within the *HOXA3* gene (cg22962123) highlighted significant hypermethylation across all three cohorts ($P = 3.11 \times 10^{-18}$).
Discussion: We present robust evidence for elevated DNA methylation associated with Alzheimer's disease neuropathology spanning the *HOXA* gene cluster on chromosome 7. These data add to the growing evidence highlighting a role for epigenetic variation in Alzheimer's disease, implicating the *HOX* gene family as a target for future investigation.

Keywords: Alzheimer's disease (AD); Braak stage; DNA methylation; Epigenetics; Epigenome-wide association study (EWAS); *HOXA*; Illumina Infinium 450K BeadChip (450K array); Meta-analysis; Neuropathology; Prefrontal cortex (PFC); Superior temporal gyrus (STG)

1. Introduction

Alzheimer's disease (AD), the most common form of dementia, is a progressive neurodegenerative disorder that is making an increasing contribution to the global burden of

[1]These authors contributed equally.
*Corresponding author. Tel.: +44 1392 408 501.
**Corresponding author. Tel.: + 44 1392 408 298.
E-mail address: j.mill@exeter.ac.uk (J.M.), k.lunnon@exeter.ac.uk (K.L.)

https://doi.org/10.1016/j.jalz.2018.01.017

disease as the population ages [1]. AD pathology is characterized by the accumulation of amyloid-β plaques and tau tangles, ultimately leading to neuronal cell loss. The neurodegeneration associated with AD is believed to start many decades before clinical onset; during this "preclinical" phase, the plaque and tangle loads in the brain increase until a person-specific threshold level is reached and behavioral changes and cognitive impairment become manifest [2–4]. At present, there are no disease-modifying treatments available, with existing medications only alleviating certain symptoms of AD. A better understanding of the underlying mechanisms precipitating the onset and progression of pathology is required to enable the design of new, more effective medications.

Increased knowledge about the functional complexity of the genome has led to speculation about the role of epigenetic variation in health and disease, including for neurodegenerative diseases such as AD [5]. Two epigenome-wide association studies (EWAS) of AD [6,7] recently identified consistent patterns of DNA methylation associated with neuropathology. Of particular interest was replicated evidence for cortex-specific hypermethylation at multiple 5′-C-phosphate-G-3′ (CpG) sites within *ANK1*, although differences at a number of other loci were identified in one or both studies [8]. One of the previously reported neuropathology-associated differentially methylated positions (DMPs), cg22962123, is located within the *HOXA* gene cluster on chromosome 7 [7]. Here, we present further evidence to support a role for altered DNA methylation in AD-associated neuropathology across an extensive region spanning the *HOXA* gene cluster.

2. Methods

2.1. Samples and subjects

Our discovery (Mount Sinai) cohort consisted of brain tissue from 147 individuals obtained from the Mount Sinai Alzheimer's Disease and Schizophrenia Brain Bank (http://icahn.mssm.edu/research/labs/neuropathology-and-brain-banking). From the 147 donors, two cortical regions (prefrontal cortex [PFC, N = 144] and superior temporal gyrus [STG, N = 142]) were used for the purposes of the study. All samples were dissected by trained specialists, snap-frozen and stored at −80°C. Further information about the samples is given in Supplementary Table 1. Ethical approval for the project was provided by the University of Exeter Medical School Research Ethics Committee under application number 14/02/041. Genomic DNA was isolated from ∼100 mg of each dissected brain region using a standard phenol-chloroform extraction protocol and tested for purity and degradation before analysis. For replication purposes, we used previously published EWAS data collected in two independent cohorts on the Illumina Infinium Human Methylation 450K BeadChip (450K array): (1) the "London" (Lunnon et al.) cohort, consisting of PFC, STG, entorhinal cortex, cerebellum (CER), and premortem blood DNA methylation data from 117 individuals from the Medical Research Council London Neurodegenerative Disease Brain Bank [6] and (2) the "Religious Orders Study/Memory and Aging Project (ROS/MAP)" (De Jager et al.) cohort, consisting of PFC DNA methylation data from 740 individuals from the Religious Orders Study and the Rush Memory and Aging Project [7]. All samples were assigned a unique code number for the experiment, which was independent of age, gender, or diagnosis. This code was used throughout the experiment and analysis.

2.2. Bisulfite treatment and Illumina Infinium BeadArray

Five hundred nanograms of genomic DNA was sodium bisulfite converted using the EZ-DNA methylation kit (Zymo Research, Orange, CA, USA), and DNA methylation was subsequently quantified using the 450K array (Illumina, USA) with arrays scanned using an Illumina iScan (software version 3.3.28). Samples were processed by tissue and randomized with respect to age and gender. The Illumina 450K array interrogates >485,000 probes covering 99% of reference sequence (RefSeq) genes, with an average of 17 CpG sites per gene region (distributed across promoter, 5′untranslated region, first exon, gene body, and 3′untranslated regions). It covers 96% of CpG islands, with additional coverage in island shores and their flanking regions.

2.3. Microarray quality control and data normalization

Initial quality control of data was conducted using GenomeStudio (version 2011.1) to determine the status of staining, extension, hybridization, target removal, sodium bisulfite conversion, specificity, and nonpolymorphic and negative controls. Probes previously reported to hybridize to multiple genomic regions or containing a single-nucleotide polymorphism at the single base extension site were removed from subsequent analyses [9,10], in addition to the 65 single-nucleotide polymorphisms used for sample identification on the array (total probes removed 72,067). For each probe, DNA methylation levels were indexed by β values, that is, the ratio of the methylated signal divided by the sum of the methylated and unmethylated signal ($M/[M + U]$).

2.4. Data analysis

All computations and statistical analyses were performed using R 3.0.2 and Bioconductor 2.13. Signal intensities were imported into R using the *methylumi* package. Initial quality control checks were performed using functions in the *methylumi* package to assess concordance between reported and genotyped gender. Non-CpG single-nucleotide polymorphism probes on the array were also used to confirm that both brain regions were sourced from the same individual where expected. Data were preprocessed and quantile normalized using the *dasen* function as part of the

wateRmelon package (wateRmelon_1.0.3) [11] within the R statistical analysis environment and batch corrected using the *ComBat* package [12]. Array data for each of the tissues were normalized separately, and initial analyses were performed separately by tissue. Full Illumina 450K array data were available for the discovery (Mount Sinai) and London (Lunnon et al.) cohorts, and thus we were able to estimate neuronal proportions in the data using the R package *CETS* [13]. For the ROS/MAP (De Jager et al.) cohort, we only had Illumina 450K array data for probes in the *HOXA* region and thus could not calculate neuronal proportions. Therefore, the effects of age, gender, and cell type composition were regressed out of the discovery (Mount Sinai) and London (Lunnon et al.) cohorts, whereas the effects of age and gender only were regressed out of the ROS/MAP (De Jager et al.) cohort before subsequent analysis. For identification of DMPs specifically altered with respect to neuropathological measures of AD, we performed a quantitative analysis in which samples were analyzed separately in each brain region using linear regression models with respect to Braak stage, with probes ranked according to *P* values. The genic location of identified DMPs was annotated by GREAT annotation [14]. We have previously established the multiple testing threshold (experiment-wide significance) for EWAS data generated on the Illumina 450K array as $P < 2.2 \times 10^{-7}$ [15]. In brief, in this previous study, 5000 permutations were performed repeating a linear regression model for randomly selected groups of cases and controls (N = 675). For each permutation, *P* values from the EWAS were saved and the minimum identified. Across all permutations, the fifth percentile was calculated to generate the 5% of α significance threshold, which was deemed to be $P < 2.2 \times 10^{-7}$. To identify differentially methylated regions (DMRs), we identified spatially correlated *P* values in our data using the Python module *comb-p* to group ≥3 spatially correlated CpGs in a 500-bp sliding window [16]. The *coMET* package was used to identify regional comethylation patterns and regional EWAS results [17]. Fisher's combined *P* value analysis was performed in the *MetaDE* package [18], and meta-analysis on correlation and case control status was performed with the *meta* package [19] within R [20]. Data are available for the discovery (Mount Sinai) cohort within Gene Expression Omnibus under accession number GSE80970. The discovery (Mount Sinai) EWAS data set has been previously used to validate the top 100 DMPs nominated in a previously published EWAS [6]. As such, we have not sought to replicate these top 100 DMPs in the present study.

3. Results

3.1. Hypermethylation associated with AD neuropathology is observed in a region spanning 48 kb across the HOXA gene cluster in the human cortex

Our primary analyses focused on matched PFC and STG tissues from 147 individuals (Supplementary Table 1). We used the 450K array to first quantify DNA methylation in the PFC and identify DMPs associated with the Braak score, a standardized measure of neurofibrillary tangle burden determined at autopsy, controlling for age, gender, and estimated neuronal cell proportion. We identified 10 experiment-wide significant ($P < 2.2 \times 10^{-7}$) DMPs (Table 1 and Fig. 1A), with 78 DMPs associated with Braak stage at a more relaxed threshold of $P < 1 \times 10^{-5}$ (Supplementary Table 2). Of these 78 DMPs, nine were located in the *HOXA* gene cluster on

Table 1
The 10 DMPs associated with Braak stage in the PFC in the discovery (Mount Sinai) cohort that reached experiment-wide significance ($P < 2.2 \times 10^{-7}$) are shown, with annotation to chromosomal location (hg19), up/downstream genes (from GREAT annotation), *P* value from our quantitative association model, and corrected DNA methylation difference (Δ) from Braak score 0–VI (as a %). Also shown is the corresponding information in the matched STG samples in the same cohort, and the matched brain regions (PFC, STG) in the London (Lunnon et al.) cohort, demonstrating a nominally significant difference. A list of the 78 top-ranked PFC DMPs at a more relaxed threshold of $P < 1 \times 10^{-5}$ is given in Supplementary Table 2

					Discovery (Mount Sinai) cohort				London (Lunnon et al.) cohort			
					Association with Braak stage				Association with Braak stage			
			GREAT annotation		PFC		STG		PFC		STG	
Probe	Location	Illumina annotation	Downstream	Upstream	Δ	*P* value	Δ	*P* value	Δ	*P* value	Δ	*P* value
cg22867816	4:16081205	PROM1	FGFBP2 (−116347)	PROM1 (+4118)	−3.90	9.80E–09	−2.04	5.21E–03	-	-	-	-
cg06977285	7:18127468		HDAC9 (−408457)	PRPS1L1 (−59983)	3.66	2.02E–08	2.68	1.84E–04	-	-	1.88	7.65E–03
cg05783384	2:218843735		RUFY4 (−90242)	TNS1 (−34885)	7.42	4.46E–08	5.55	8.01E–05	3.26	7.76E–03	3.83	6.48E–04
cg07349815	3:123751269		CCDC14 (−70706)	KALRN (−62258)	5.15	6.70E–08	-	-	2.15	.02	1.83	7.35E–03
cg21806242	11:72532891	ATG16L2	ATG16L2 (+7539)	FCHSD2 (+320414)	8.51	7.02E–08	5.55	4.08E–04	5.22	3.86E–04	4.62	1.10E–03
cg03834767	7:90794392	CDK14	FZD1 (−99390)	CDK14 (+455681)	−4.50	8.13E–08	-	-	-	-	-	-
cg13935577	12:107974897	BTBD11	PWP1 (−104611)	BTBD11 (+262708)	9.11	8.45E–08	5.27	1.49E–03	4.02	5.10E–03	3.73	.02
cg27078890	11:128457459	ETS1	ETS1 (−23)		4.85	9.86E–08	-	-	2.09	.02	-	-
cg22962123	7:27153605	HOXA3	HOXA2 (−11176)	HOXA3 (+5608)	7.88	1.20E–07	5.12	2.78E–04	5.62	2.24E–05	5.18	5.21E–04
cg26199857	12:54764265	ZNF385A	GPR84 (−5995)	ZNF385A (+20,816)	5.43	1.87E–07	4.44	1.02E–03	2.62	.03	-	-

Abbreviations: DMP, differentially methylated position; PFC, prefrontal cortex; STG, superior temporal gyrus.

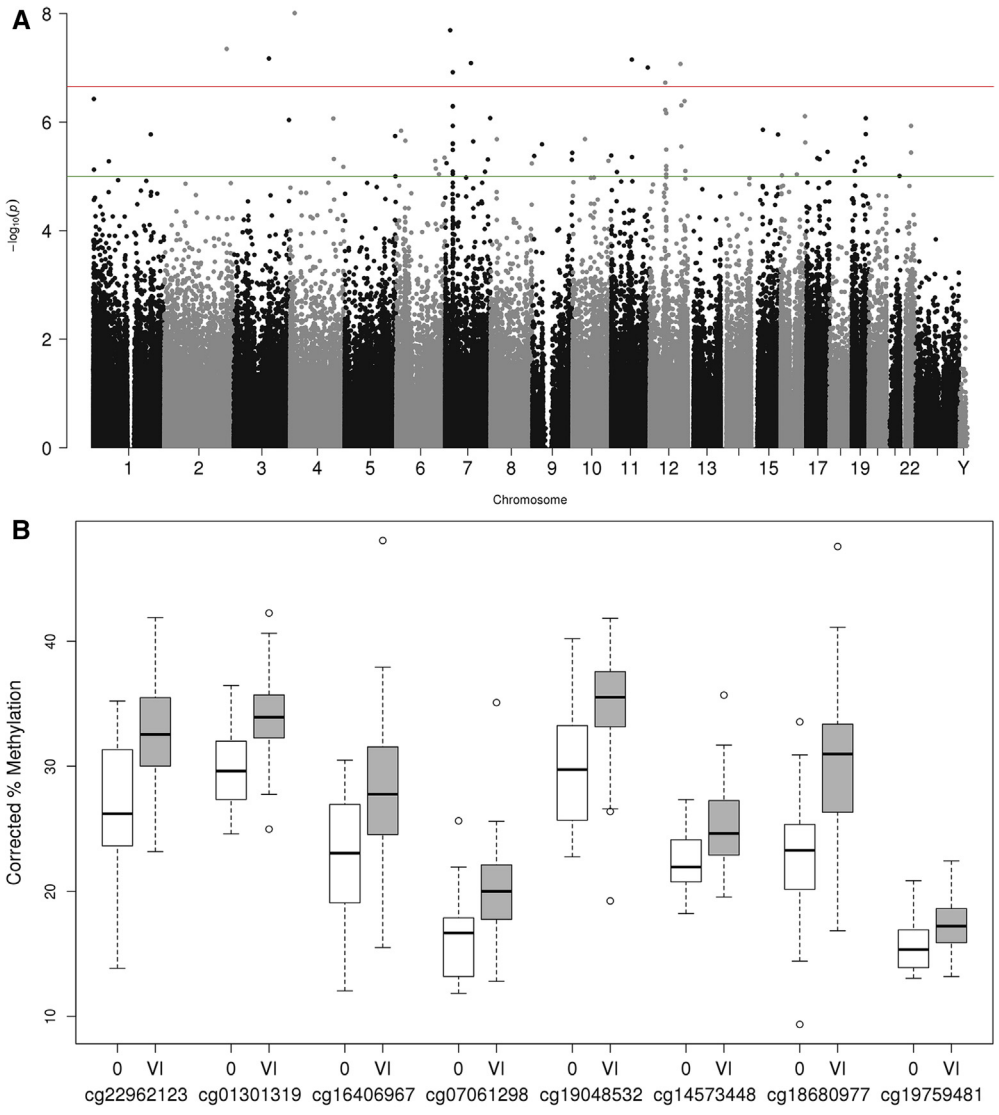

Fig. 1. *HOXA3* hypermethylation is associated with neuropathological measures of AD in cortex. (A) A Manhattan plot of association between DNA methylation in the PFC with Braak stage highlights associations at loci across the genome, with a region on chromosome 7 in the *HOXA3* gene showing the greatest number of probes associated with pathology. The red line indicates experiment-wide significance threshold ($P = 2.2 \times 10^{-7}$), with the green line indicating a more relaxed significance threshold ($P = 1 \times 10^{-5}$). (B) Using a sliding window approach to identify differentially methylated regions, we identified six within the *HOXA* gene cluster (Table 2), with the most significant region spanning 364 bp in the *HOXA3* gene and containing seven CpG sites that showed increased methylation in late-stage AD (Braak stage VI) compared to healthy controls (Braak stage 0). (C) A mini-Manhattan plot across the *HOXA* gene cluster. Highlighted between gray dashed lines is a 48,754-bp region containing 208 probes characterized by neuropathology-associated hypermethylation. Red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI), and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. (D) The site demonstrating the greatest DNA methylation difference (cg22962123) in the PFC (R = 0.36, $P = 1.2 \times 10^{-7}$) also showed a similar but weaker association in the STG (R = 0.28, $P = 2.78 \times 10^{-4}$). (E) A quadrant plot of the effect size of the 208 probes identified in the PFC and their corresponding effect size in the STG highlights a significant correlation between brain regions (R = 0.76, $P = 2.66 \times 10^{-40}$). Abbreviations: AD, Alzheimer's disease; CpG, 5′-C-phosphate-G-3′; PFC, prefrontal cortex; STG, superior temporal gyrus.

chromosome 7, most notably in the vicinity of *HOXA3*, with one *HOXA* DMP reaching experiment-wide significance (cg22962123: $P = 1.2 \times 10^{-7}$). We next used a sliding window approach (*comb-p* [16]) to identify spatially correlated regions of differential DNA methylation associated with neuropathology; Table 2 lists DMRs spanning at least three probes with a window size of 500 bp and a Sidak-corrected *P* value < .05. We identified six closely located DMRs within the *HOXA* gene region, with the most significant DMR in the *HOXA* region spanning seven probes in a 364-bp region within intron 1 of *HOXA3* (Fig. 1B; Sidak-corrected $P = 1.19 \times 10^{-9}$). Of note, we observed an extended region of neuropathology-associated hypermethylation spanning 48,754 bp from upstream of the *HOXA2* gene to the *HOXA6*

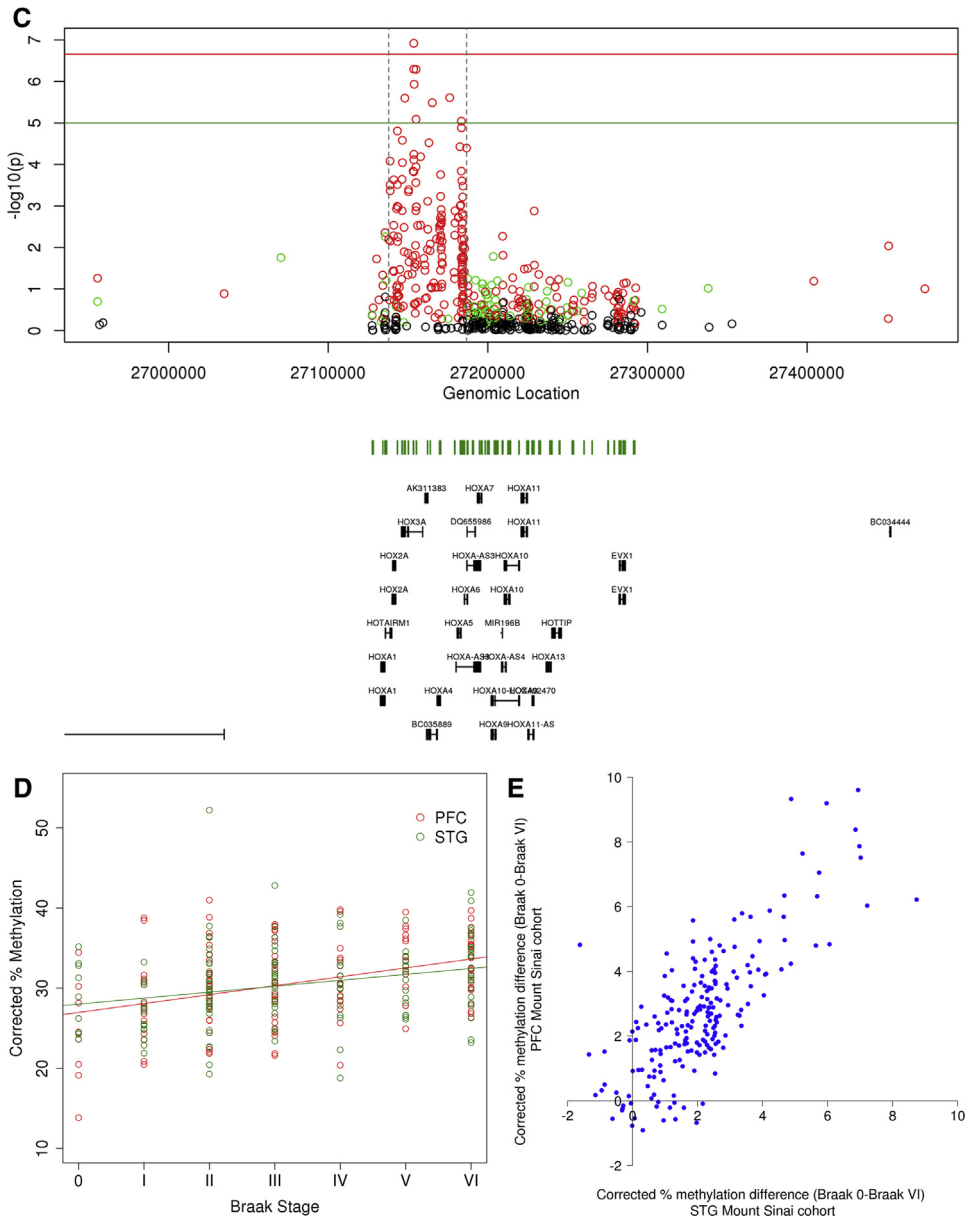

Fig. 1. *Continued*

gene and covering 208 Illumina 450K array probes (Fig. 1C). Given that DNA methylation at nearby CpG sites can be highly correlated [21], we visualized comethylation patterns between CpG sites within *HOXA3* using *coMET* [17] and observed highly correlated patterns of DNA methylation between CpG sites in this extended region (Supplementary Fig. 1). We next sought to test whether neuropathology-associated DNA methylation patterns across this 48,754-bp region were specific to the PFC, using the Illumina 450K array to profile STG samples from the same individuals. In total, seven probes in the region demonstrated significantly increased DNA methylation after correcting for 208 tests ($P < 2.4 \times 10^{-4}$), with the top PFC DMP (cg22962123) being similarly hypermethylated with respect to Braak stage (Fig. 1D; PFC: R = 0.36, $P = 1.2 \times 10^{-7}$; STG: R = 0.28, $P = 2.78 \times 10^{-4}$). There was an overall consistent pattern of effect sizes across both brain regions for the 208 probes in the *HOXA* neuropathology-associated region (Fig. 1E; R = 0.76, $P = 2.66 \times 10^{-40}$).

3.2. Cortical neuropathology-associated hypermethylation in HOXA3 is observed in independent study cohorts

We next sought to replicate the observation of neuropathology-associated hypermethylation across these 208 probes in two independent, previously published data sets. First, we examined the "London" (Lunnon et al. [6])

Table 2
DMRs associated with Braak stage in the PFC—Shown are all significantly associated regions (Sidak-corrected *P* value < .05) that contain three or more probes, with chromosomal location (hg19), up/downstream genes, number of probes in the significant region, and Sidak-corrected *P* value

Chr	Start	End	Illumina annotation	GREAT annotation		Number of probes	Šidák-corrected *P* value
Chr11	2,321,770	2,323,247	C11ORF21	TSPAN32 (−734)	C11orf21 (+634)	27	3.20E–11
Chr7	27,153,580	27,153,944	HOXA3	HOXA2 (−11332)	HOXA3 (+5452)	7	1.19E–09
Chr7	27,154,262	27,155,234	HOXA3	HOXA2 (−12318)	HOXA3 (+4466)	16	4.31E–09
Chr7	27,169,957	27,171,401	HOXA4	HOXA4 (−261)		21	2.13E–08
Chr11	3,15,908	3,16,456	IFITM1 Closest	IFITM1 (+2329)	IFITM3 (+4868)	5	4.02E–08
Chr12	58,119,915	58,120,237	AGAP2	AGAP2 (+11,953)	OS9 (+32,172)	6	1.22E–07
Chr7	27,183,133	27,184,853	HOXA5/HOXA-AS3	HOXA5 (−706)		42	2.19E–06
Chr5	78,985,425	78,985,900	CMYA5	CMYA5 (−37)		10	2.31E–06
Chr19	10,736,006	10,736,448	SLC44A2	SLC44A2 (+293)		8	3.68E–06
Chr19	39,086,733	39,087,186	MAP4K1	MAP4K1 (+21,604)	RYR1 (+162490)	4	4.94E–06
Chr6	10,556,147	10,556,523	GCNT2	GCNT6 (−77,658)	GCNT2 (+27,746)	3	2.93E–05
Chr3	194,014,592	194,015,171	GRM2 Closest	CPN2 (+57,175)	HES1 (+160,948)	4	3.24E–05
Chr4	184,908,351	184,909,018	STOX2	STOX2 (+82,176)	ENPP6 (+230,429)	8	3.60E–05
Chr7	27,145,972	27,146,445	HOXA3	HOXA2 (−3779)		5	4.11E–05
Chr17	46,388,390	46,388,465	SKAP1	SKAP1 (+119,124)	SNX11 (+203,508)	3	4.77E–05
Chr17	74,475,240	74,475,402	RHBDF2	RHBDF2 (+22,168)	AANAT (+25,888)	5	8.13E–05
Chr3	51,740,741	51,741,280	GRM2	GRM2 (−75)		6	1.93E–04
Chr17	41,363,502	41,364,121	NBR1/TMEM106A	TMEM106 A (−82)		11	3.04E–04
Chr17	43,318,610	43,319,371	FMNL1	FMNL1 (+19,835)	SPATA32 (+20,488)	6	4.51E–04
Chr7	158,281,410	158,281,613	PTPRN2	PTPRN2 (+98,859)		3	4.66E–04
Chr13	43,565,901	43,566,496	EPSTI1	DNAJC15 (−31,140)	TNFSF11 (+417,910)	9	4.72E–04
Chr20	57,582,787	57,583,520	CTSZ Closest	CTSZ (−852)		18	6.82E–04
Chr19	3,179,545	3,180,035	S1PR4	NCLN (−5808)	S1PR4 (+1054)	4	7.59E–04
Chr22	37,608,611	37,608,819	SSTR3 Closest	SSTR3 (−353)		3	8.84E–04
Chr13	113,698,408	113,699,016	MCF2L	F7 (−61,409)	MCF2L (+75,177)	13	9.15E–04
Chr9	34,457,129	34,457,500	FAM219A	DNAI1 (−1518)		4	1.05E–03
Chr17	75,315,081	75,315,567	SEPT9	TNRC6C (−685,813)	SEPT9 (+37,832)	8	1.28E–03
Chr16	29,674,618	29,675,214	SPN	SPN (+336)		6	1.77E–03
Chr1	55,246,867	55,247,408	TTC22	PARS2 (−16,951)	DHCR24 (+105,753)	5	2.45E–03
Chr12	58,132,558	58,133,008		AGAP2 (−754)		3	3.00E–03
Chr7	27,138,712	27,138,974	HOTAIRM1	HOXA1 (−3250)		4	3.19E–03
Chr16	67,686,832	67,687,392	RLTRP	ACD (+7534)	RLTPR (+8290)	4	3.59E–03
Chr12	58,129,855	58,130,410	AGAP2	AGAP2 (+1896)	OS9 (+42,229)	4	4.42E–03
Chr17	19,314,299	19,314,618	RNF112	RNF112 (−48)		6	9.80E–03
Chr15	40,583,227	40,583,422	PLCB2	PLCB2 (+16,798)	PAK6 (+51,704)	3	.01922
Chr15	38,988,533	38,988,860	C15ORF53	THBS1 (−884,597)	RASGRP1 (−131,690)	4	.01974
Chr16	1,482,952	1,483,192	CCDC154 Closest	C16orf91 (−3727)		3	.02843

Abbreviations: DMR, differentially methylated region; PFC, prefrontal cortex.

data set, comprising Illumina 450K array data generated using matched PFC, STG, entorhinal cortex, CER, and premortem blood samples obtained from 117 donors (described in [6]; Supplementary Table 1). We observed a similar pattern of Braak-associated DNA methylation across this 208-probe region in the replication cohort in both the PFC (Fig. 2A) and STG (Supplementary Fig. 2), with a highly correlated effect size between cohorts in both brain regions (PFC: Fig. 2B; R = 0.74, $P = 2.27 \times 10^{-37}$; STG: Supplementary Fig. 3; R = 0.68, $P = 1.87 \times 10^{-29}$)—15 probes in the PFC and 6 probes in the STG reaching our corrected significance threshold ($P < 2.4 \times 10^{-4}$). In contrast, no probes in this region reached the corrected significance threshold in the entorhinal cortex (Supplementary Fig. 4), although the effect size was still correlated (R = 0.41, $P = 1.23 \times 10^{-9}$). Similarly, no probes reached the significance threshold in the CER (Supplementary Fig. 5) or in premortem whole blood collected in a subset (N = 57) of the same individuals (Supplementary Fig. 6), with no correlation of effect sizes in either the CER (R = 0.03, *P* = .639) or blood (R = 0.11, *P* = .138). This indicates that the association may be specific to only particular regions of the cortex.

We subsequently assessed this region in the "ROS/MAP" (De Jager et al.) data set comprising of 740 PFC samples profiled on the Illumina 450K array (as described in De Jager et al. [7]; Supplementary Table 1) observing a similar pattern of effects with highly significant neuropathology-associated hypermethylation across probes in the *HOXA* genic region (Fig. 2C), and a significant correlation of effect size with the same 208 probes in the PFC in the discovery cohort (Fig. 2D; R = 0.80, $P = 2.39 \times 10^{-48}$). A Fisher's combined

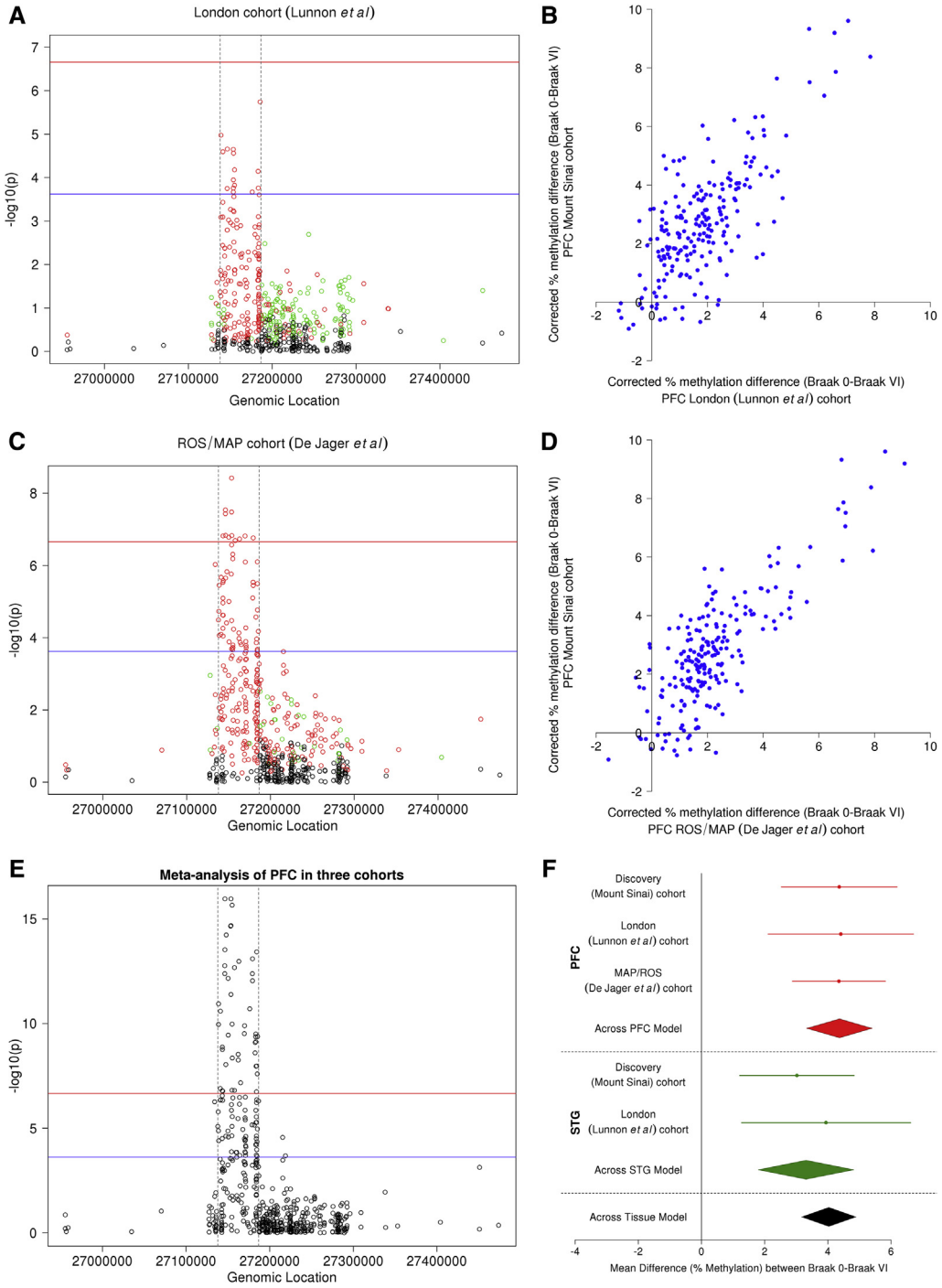

Fig. 2. Replication of neuropathology-associated DNA methylation differences across the *HOXA* gene cluster in additional study cohorts. (A) We identified a consistent pattern of increased DNA methylation across the *HOXA* cluster in the London (Lunnon et al.) cohort in the PFC (B) with a strong correlation in effect size across the 208 probes in the region between data sets (R = 0.74, $P = 2.24 \times 10^{-37}$). (C) A similar pattern of DNA methylation changes was observed in the PFC in the ROS/MAP (De Jager et al.) cohort, (D) with a strong correlation in effect size across the 208 probes in the region between data sets (R = 0.80, $P = 2.39 \times 10^{-48}$). (E) A Fisher's combined *P* value meta-analysis of the PFC with respect to Braak stage across all three cohorts showed striking patterns of increased DNA methylation with many probes in the *HOXA3* region reaching experiment-wide significance. (F) The most significant probe identified from the discovery cohort (cg22962123) was also the most significant probe in the meta-analysis ($P = 3.11 \times 10^{-18}$) and characterized by neuropathology-associated hypermethylation across all three cohorts. In plots (A) and (C) red circles indicate increased DNA methylation in disease (≥1% between Braak 0 and Braak VI), green circles indicate decreased DNA methylation in disease (≥1% between Braak 0 and Braak VI), and black circles indicate DNA methylation differences <1% between Braak 0 and Braak VI. In plots (A), (C), and (E), the red line indicates experiment-wide significance ($P = 2.2 \times 10^{-7}$), whereas the blue line indicates significance after correcting for 208 tests ($P = 2.4 \times 10^{-4}$). In plots (A), (C), and (E), the gray dashed lines indicate the same region across the graphs for reference. These show the probes in the extended DMR. In plot (F), red denotes the PFC and green denotes the STG. Abbreviations: DMR, differentially methylated region; PFC, prefrontal cortex; ROS/MAP, Religious Orders Study/Memory and Aging Project; STG, superior temporal gyrus.

P value of DNA methylation differences across this region in all three PFC data sets confirmed a clearly defined region of significant neuropathology-associated elevated DNA methylation, with many individual DMPs passing the threshold for experiment-wide significance (Fig. 2E), and a consistent pattern of effects across the three cohorts (Supplementary Fig. 7). The most significant DMP identified within the *HOXA3* gene in our discovery cohort (cg22962123; Table 1) was also the most significant DMP in our Fisher's combined *P* value analysis ($P = 1 \times 10^{-20}$). A meta-analysis comparing Braak 0 to VI demonstrated increased DNA methylation with respect to Braak stage across all cohorts in the PFC (Fig. 2F; $P = 3.11 \times 10^{-18}$). Together, our data suggest that DNA hypermethylation across the extended *HOXA* gene region is robustly associated with AD-related neuropathology in both the PFC and STG, with the strongest effects in the vicinity of *HOXA3*.

4. Discussion

We identified an extended region of elevated DNA methylation in the *HOXA* gene cluster that is associated with AD neuropathology, with consistent effects seen across three independent postmortem brain sample cohorts. Although one previous study had demonstrated differential methylation at a single CpG within the *HOXA* gene cluster [7] and another identified a DMR spanning seven CpG sites [6], this represents the first study to illustrate that hypermethylation in this region extends to 208 DMPs, spanning approximately 48.7 Kb. Differential DNA methylation in the *HOXA* gene cluster has been previously reported in blood collected from Down syndrome individuals [22], which is interesting given that many Down syndrome individuals develop AD resulting from an additional copy of the *APP* gene due to trisomy on chromosome 21. The Down syndrome study demonstrated differential DNA methylation in 20 probes largely located within *HOXA2*. Of note, 17 of these probes were significantly hypermethylated in the PFC in our discovery (Mount Sinai) cohort. However, none were differentially methylated in premortem blood in the London (Lunnon et al.) cohort. In the context of other neurodegenerative disorders, one study that investigated microRNAs targeting *HOX* genes in Huntington's disease demonstrated increased levels of microRNAs related to *HOXA5*, *HOXA10*, *HOXA11*, *HOXA-11AS*, *HOXA13*, and *HOTAIRM1* in the PFC in Huntington's disease [23]. Although *HOX* genes encode potent transcription factors that play a critical role in embryonic development [24], a recent study in *Drosophila* also highlighted a potent protective function for *HOX* genes in neurons, implicating a role in neuroprotection [25]. Interestingly, this study also highlighted how *HOX* genes act to maintain expression of the ankyrin locus, an important observation given our previous finding of altered DNA methylation in *ANK1* in AD [6]. Indeed, to further explore this hypothesis, we examined the correlation between DNA methylation levels at the most significant *HOX* probe identified in the present study (cg22962123) with the two *ANK1* DMPs that we previously identified to be associated with AD neuropathology (cg11823178 and cg05066959) [6,7] in the PFC, identifying a significant correlation with both *ANK1* probes (cg11823178: $R = 0.24$, $P = 5.15 \times 10^{-10}$; cg05066959: $R = 0.20$, $P = 2.93 \times 10^{-8}$). Although this correlation could reflect the association between both *HOXA3* and *ANK1* probes with Braak stage, it could highlight a novel physiological mechanism, particularly as we still observed significant hypermethylation ($P = 1.67 \times 10^{-5}$) at our top *HOXA* probe (cg22962123), when controlling for levels of DNA methylation in the top *ANK1* probe (cg11823178). Looking to the future, analyses of gene expression levels should be performed to facilitate the interpretation of the DNA methylation differences we observe in *HOXA*. To conclude, this study provides further evidence for altered epigenetic processes in the pathophysiology of AD and suggests that further work on the neuroprotective functions of *HOX* genes is warranted.

Acknowledgments

This work was funded by NIH grant R01 AG036039 to J.M. and by Alzheimer's Society grant AS-PG-14-038, Alzheimer's Association grant NIRG-14-320878, and a grant from BRACE (Bristol Research into Alzheimer's and Care of the Elderly) to K.L. Brain banking and neuropathology assessments for the discovery cohort from the Mount Sinai Alzheimer's disease and Schizophrenia Brain Bank were supported by NIH grants AG02219, AG05138, and MH064673 and the Department of Veterans Affairs VISN3 MIRECC. Brain banking and neuropathological assessment for the London (Lunnon et al.) cohort was provided by The London Neurodegenerative Diseases Brain Bank, which receives funding from the Medical Research Council (MRC) and as part of the Brains for Dementia Research (BDR) programme, jointly funded by Alzheimer's Research UK and Alzheimer's Society. The ROS/MAP (De Jager et al.) cohort was supported by the National Institutes of Health grants: R01 AG036042, R01AG036836, R01 AG17917, R01 AG15819, R01 AG032990, R01 AG18023, RC2 AG036547, P30 AG10161, P50 AG016574, U01 ES017155, KL2 RR024151, and K25 AG041906-01.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jalz.2018.01.017.

RESEARCH IN CONTEXT

1. Systematic review: We performed an epigenome-wide association study (EWAS) to identify differential DNA methylation associated with Braak stage in a discovery cohort of 147 individuals. A regional analysis identified six differentially methylated regions (DMRs), consisting of >3 differentially methylated positions (DMPs) with a Sidak-corrected *P* value <0.05, within the *HOXA* gene cluster. Further investigation highlighted a region of neuropathology-associated hypermethylation spanning >48kb (208 probes) across the *HOXA* gene cluster.
2. Interpretation: *HOX* genes encode transcription factors important in neural development. A recent study has provided evidence that *Hox* genes can maintain expression of the *Ank* locus [25], which is particularly interesting given that two previous EWAS have provided robust evidence for differential DNA methylation in AD cortex in the *ANK1* gene [6,7]. A significant correlation of DNA methylation was seen between the most significant *HOX* probe identified in the current study with the two *ANK1* DMPs previously identified [6,7], even when controlling for levels of DNA methylation in *ANK1*.
3. Future directions: Analyses of gene expression levels should be performed to facilitate the interpretation of the DNA methylation differences we observed in *HOXA* as well as determining whether these DNA methylation changes are a cause or consequence of AD neuropathology.

References

[1] Prince M, Guerchet M, Prina M. The Global Impact of Dementia 2013–2050. London, United Kingdom: Alzheimer's Disease International (ADI); 2013.

[2] Blennow K, de Leon MJ, Zetterberg H. Alzheimer's disease. Lancet 2006;368:387–403.

[3] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011;7:280–92.

[4] Jack CR Jr, Knopman DS, Jagust WJ, Shaw LM, Aisen PS, Weiner MW, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. Lancet Neurol 2010;9:119–28.

[5] Lunnon K, Mill J. Epigenetic studies in Alzheimer's disease: current findings, caveats, and considerations for future studies. Am J Med Genet B Neuropsychiatr Genet 2013;162B:789–99.

[6] Lunnon K, Smith R, Hannon E, De Jager PL, Srivastava G, Volta M, et al. Methylomic profiling implicates cortical deregulation of ANK1 in Alzheimer's disease. Nat Neurosci 2014; 17:1164–70.

[7] De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nature Neurosci 2014; 17:1156–63.

[8] Lord J, Cruchaga C. The epigenetic landscape of Alzheimer's disease. Nat Neurosci 2014;17:1138–40.

[9] Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 2013;8:203–9.

[10] Price ME, Cotton AM, Lam LL, Farre P, Emberly E, Brown CJ, et al. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin 2013;6:4.

[11] Pidsley R, Wong CCY, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC genomics 2013;14:293.

[12] Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 2012;28:882–3.

[13] Guintivano J, Aryee M, Kaminsky Z. A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. Epigenetics 2013;8:290–302.

[14] McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 2010;28:495–501.

[15] Hannon E, Dempster E, Viana J, Burrage J, Smith AR, Macdonald R, et al. An integrated genetic-epigenetic analysis of schizophrenia: Evidence for co-localization of genetic associations and differential DNA methylation. Genome Biol 2016;17:176.

[16] Pedersen BS, Schwartz DA, Yang IV, Kechris KJ. Comb-p: software for combining, analyzing, grouping and correcting spatially correlated P-values. Bioinformatics 2012;28:2986–8.

[17] Martin TC, Yet I, Tsai PC, Bell JT. coMET: visualisation of regional epigenome-wide association scan results and DNA co-methylation patterns. BMC bioinformatics 2015;16:131.

[18] Wang X, Li J, Tseng GC. MetaDE: Microarray meta-analysis for differentially expressed gene detection; 2012.

[19] Schwarzer G, Carpenter JR, Rücker G. Meta-Analysis with R. Cham, Switzerland: Springer; 2015. https://doi.org/10.1007/978-3-319-21416-0.

[20] Wang X, Kang DD, Shen K, Song C, Lu S, Chang LC, et al. An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. Bioinformatics 2012;28:2534–6.

[21] Bell JT, Pai AA, Pickrell JK, Gaffney DJ, Pique-Regi R, Degner JF, et al. DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines. Genome Biol 2011; 12:R10.

[22] Bacalini MG, Gentilini D, Boattini A, Giampieri E, Pirazzini C, Giuliani C, et al. Identification of a DNA methylation signature in blood cells from persons with Down Syndrome. Aging 2015;7:82–96.

[23] Hoss AG, Kartha VK, Dong X, Latourelle JC, Dumitriu A, Hadzi TC, et al. MicroRNAs located in the Hox gene clusters are implicated in huntington's disease pathogenesis. PLoS Genet 2014;10:e1004188.

[24] Krumlauf R. Hox genes in vertebrate development. Cell 1994; 78:191–201.

[25] Friedrich J, Sorge S, Bujupi F, Eichenlaub MP, Schulz NG, Wittbrodt J, et al. Hox function is required for the development and maintenance of the Drosophila feeding motor unit. Cell Rep 2016;14:850–60.